

ABSTRACT OF THE DISCLOSURE

Systems and methods are described that facilitate predictive web-crawling in a computer environment. Aspects of the invention provide for predictive, utility-based, and decision theoretic probability assessments of changes in subsets of web pages, enhancing web-crawling ability and ensuring that web page information is maintained in a fresh state. Additionally, the invention facilitates selective crawling of pages with a high probability of change.